THE ROYAL SOCIETY

# Analyses of mitochondrial genomes strongly support a hippopotamus–whale clade

**Björn M. Ursing*** **and Ulfur Arnason**

*Department of Genetics, Division of Evolutionary Molecular Systematics, Lund University, Sölvegatan 29, S-223 62 Lund, Sweden*

Although the sister-group relationship between Cetacea and Artiodactyla is widely accepted, the actual artiodactyl group which is closest to Cetacea has not been conclusively identified. In the present study, we have sequenced the complete mitochondrial genome of the hippopotamus, *Hippopotamus amphibius*, and included it in phylogenetic analyses together with 15 other placental mammals. These analyses separated the hippopotamus from the other suiform included, the pig, and identified the hippopotamus as the artiodactyl sister group of the cetaceans, thereby making both Artiodactyla and the suborder Suiformes paraphyletic. The divergence between the hippopotamid and cetacean lineages was calculated using this molecular data and was estimated at *ca.* 54 Ma BP.

**Keywords:** molecular phylogeny; mitochondrial DNA; Artiodactyla; *Hippopotamus amphibius*; Suiformes; Cetacea

## 1. INTRODUCTION

Artiodactyla is a diverse order with about 80 living genera and a rich fossil record dating from the late Palaeocene–early Eocene. The artiodactyls are a well-defined group, with the most pronounced external characteristic being the paraxonic foot, the axis of which passes between the third and fourth digits, irrespective of the number of digits retained. The first digits are almost always absent, even in the most primitive artiodactyls, and among the most specialized groups the second and fifth digits are also rudimentary or absent. Living artiodactyls may be grouped into eight distinct families: pigs, peccaries, hippopotamuses, bovids, deer, tragulids, giraffes and camels. The higher classification within the order Artiodactyla is problematic, but most schemes of classification divide extant artiodactyls into three suborders: Suiformes (pigs, peccaries and hippopotamuses), Ruminantia (bovids, deer, tragulids and giraffes), and Tylopoda (camels), for example, see Colbert & Morales (1991).

The first character state phylogenetic analysis of a complete mitochondrial (mt) gene, cytochrome *b*, identified artiodactyls and cetaceans as sister groups (Irwin *et al.* 1991). Extended analyses of the same gene, including the hippopotamus, placed cetacean origin within the Artiodactyla, reconstructing the relationship (outgroup (Suidae (Ruminantia (Hippopotamidae, Cetacea)))) and thereby making Artiodactyla as well as Suiformes paraphyletic (Irwin & Arnason 1994; Arnason & Gullberg 1996). Analyses of nuclear data have also identified a sister-group relationship between Cetacea and Hippopotamidae (Sarich 1993; Gatesy *et al.* 1996; Gatesy 1997), although the Ruminantia–Hippopotamidae–Cetacea relationship remained unresolved in analyses based on the positions of SINE (short interspersed elements) insertion sites (Shimamura *et al.* 1997).

These molecular findings, suggesting artiodactyl as well as suiform paraphyly, have made it necessary to examine both artiodactyl and artiodactyl–cetacean relationships on the basis of extensive data sets such as complete mtDNA sequences. Hitherto, three complete artiodactyl mtDNAs, those of the cow (Anderson *et al.* 1982), the pig (Ursing & Arnason 1998), and the sheep (Hiendleder *et al.* 1998), have been reported. These sequences together with the presently described complete mtDNA of the hippopotamus, *Hippopotamus amphibius*, and two cetacean mtDNAs, those of the fin (Arnason *et al.* 1991) and the blue (Arnason & Gullberg 1993) whales, allow central questions of artiodactyl and artiodactyl–cetacean relationships to be addressed using a more extensive molecular data set.

## 2. METHODS

Isolated mtDNA from the hippopotamus was provided by Dr Eric Harley, Department of Chemical Pathology, University of Cape Town, South Africa. This mtDNA was digested separately with *Spe* I and *Bln* I. The resulting fragments were ligated directly into the phage vector M13mp18 and/or mp19. Regions not covered by natural clones were PCR-amplified then ligated as above. Sequencing was performed manually on single-stranded DNA, using the dideoxy method with $^{35}$S-dATP (Sanger 1981) and both the universal and numerous specific oligodeoxynucleotide primers. Part of the *NADH2* gene was sequenced directly (DNA sequencer model 4000L, LI-COR Inc.) using the Thermo Sequenase cycle-sequencing kit (Amersham) with 7-deaza-dGTP and several fluorescent-labelled primers.

Phylogenetic analyses were carried out on the concatenated sequences of the 12 protein-coding genes encoded by the mitochondrial H-strand. The *NADH6* gene was not included as the

*Author for correspondence (bjorn.ursing@gen.lu.se).


*Received* 7 August 1998 *Accepted* 3 September 1998

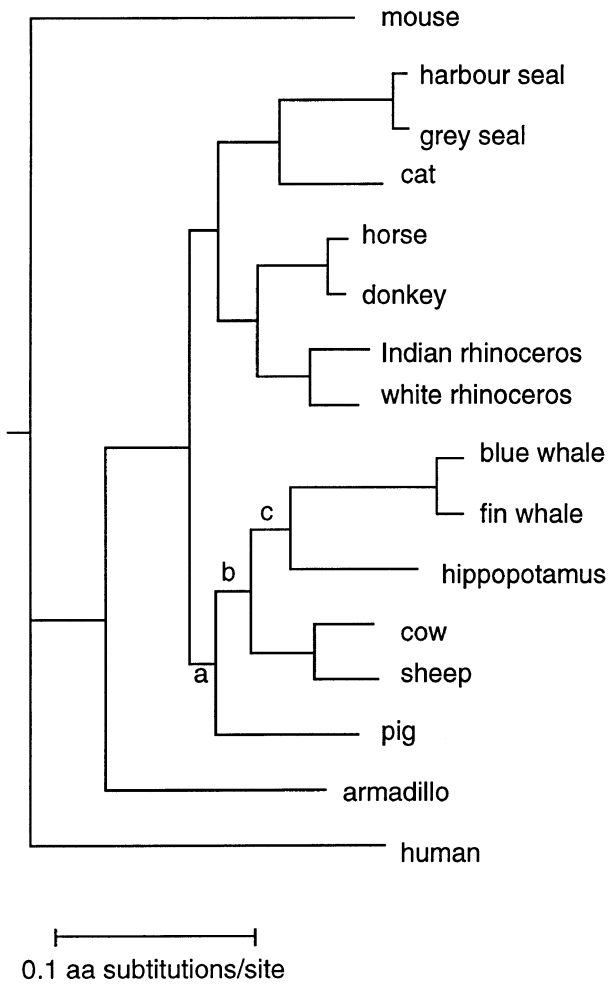


Figure 1. Maximum-likelihood tree reconstructed by PUZZLE version 4.0 (Strimmer & von Haeseler 1996), using the mtREV-24 model for aa sequence evolution (Adachi & Hasegawa 1996*b*), and showing the position of the hippopotamus relative to that of a number of other mammalian species. The tree is based on concatenated aa sequences of 12 mitochondrial protein-coding genes. After excluding gaps and ambiguous sites adjacent to gaps, the length of the alignment was 3549 aa. The branch lengths are proportional to the genetic distances between the taxa. The support values for the branches labelled a–c are given in table 1.

nucleotide (nt) composition of this gene deviates from that of the other mitochondrial protein-coding genes, thereby violating the assumptions of some of the phylogenetic algorithms used. In order to ensure the comparison of homologous sites, stop codons, gaps, and ambiguous sites adjacent to gaps, were excluded. The length of the alignment was 10 647 nt or 3549 amino acids (aa). The rationale for preferring phylogenetic analysis of large data sets, such as the concatenated protein-coding sequences of complete mtDNA genomes, is that the stochastic effects occurring with small data sets are gradually reduced with increasing alignment length (Cao *et al.* 1994).

The analyses included the following species: mouse, *Mus musculus* (Bibb *et al.* 1981); harbour seal, *Phoca vitulina* (Arnason & Johnsson 1992); grey seal, *Halichoerus grypus* (Arnason *et al.* 1993*a*); cat, *Felis catus* (Lopez *et al.* 1996); horse, *Equus callabus* (Xu & Arnason 1994); donkey, *Equus asinus* (Xu *et al.* 1996*a*); Indian rhinoceros, *Rhinoceros unicornis* (Xu *et al.* 1996*b*); white rhinoceros, *Ceratotherium simum* (Xu & Arnason 1997); blue whale, *Balaenoptera musculus* (Arnason & Gullberg 1993); fin whale, *Balaenoptera physalus* (Arnason *et al.* 1991); hippopotamus, *Hippopotamus amphibius* (present study); cow, *Bos taurus* (Anderson *et al.* 1982); sheep, *Ovis aries* (Hiendleder *et al.* 1998); pig, *Sus scrofa* (Ursing & Arnason 1998); armadillo, *Dasypus novemcinctus* (Arnason *et al.* 1997); human ('Lund'), *Homo sapiens* (Arnason *et al.* 1996).

Table 1. *Support values for specific branches*, a–c, *shown in figure 1*

(Support values for QP–ML were established using 1000 puzzling steps. For NJ and MP, bootstrap values were based on 1000 replicates on the nt level and 100 replicates on the aa level.)

| method | alignment | a | b | c |
|---|---|---|---|---|
| QP–ML | aa | 84 | 82 | 97 |
| | nt | 89 | 54 | 99 |
| MP | aa | 55 | 75 | 82 |
| | nt | 34 | 98 | 100 |
| NJ | aa | 87 | 97 | 98 |
| | nt | 98 | 100 | 100 |

Analyses were carried out at both the aa and the nt levels, applying three different approaches to phylogenetic reconstruction: maximum likelihood, ML (Felsenstein 1991); neighbour joining, NJ (Saitou & Nei 1987); and maximum parsimony, MP (Fitch 1971); implemented by the PUZZLE, version 4.0 (Strimmer & von Haeseler 1996), the PHYLIP, version 3.52c (Felsenstein 1991), or the MOLPHY, version 2.2 (Adachi & Hasegawa 1996*a*) program packages. The support values for ML were calculated using quartet puzzling (QP), QP–ML.

The nt analyses included all non-synonymous changes at the first codon position, all second codon position changes, and third codon position transversions. At the nt level, the HKY (Hasegawa, Kishino and Yano) model for sequence evolution (Hasegawa *et al.* 1985) was used with 1000 bootstrap replicates, and at the aa level the Dayhoff matrix was used (Dayhoff 1978). Due to computational constraints, the analyses at the aa level were limited to 100 bootstrap replicates. In the case of QP–ML the analyses were made with 1000 puzzling steps, the TN model for nt sequence evolution (Tamura & Nei 1993), and the mtREV-24 model for aa sequence evolution (Adachi & Hasegawa 1996*b*).

The dating of the hippopotamus–whale split was estimated using QP–ML aa distances. The rate for the hippopotamus was used for the internal branch leading to the whale–hippopotamus split (branch c in figure 1).

The hippopotamus mtDNA sequence has been deposited at the EMBL database with accession number AJ010957. Users of the sequence are kindly requested to refer to the present publication and not only to the accession number of the sequence.

## 3. RESULTS

The length of the hippopotamus mtDNA is 16 407 nt. The organization of the molecule conforms with that of other complete mammalian mtDNAs. The nt and aa composition of the individual genes is similar to that of other artiodactyls and to the fin and blue whales. All of the hippopotamus protein-coding genes have a methionine (ATG or ATA) initiation codon, except for *NADH3* which has an isoleucine (ATT) start codon and *NADH4L*

Table 2. *Phylogenetic analysis of topologies of the artiodactyl–cetacean clade*

(The ML analysis was based on the mtREV-24 model of sequence evolution. The value in angled brackets shows the log-likelihood (ln L) value of the best tree. Δln L indicates the difference in ln L to that of the best tree, followed by the standard error (s.e.; Kishino & Hasegawa 1989) and bootstrap probability (pBoot), for this particular topology (Kishino *et al*. 1990). OUT: outgroup; SUI: pig; RUM: ruminants (cow, sheep); HIPPO: hippopotamus; CET: whales (blue whale, fin whale). The ln L, s.e., and pBoot values were calculated using NucML and ProtML in the MOLPHY, version 2.2 (Adachi & Hasegawa 1996*a*) program.

| topology | amino acid | | | nucleotide | | |
|---|---|---|---|---|---|---|
| | ΔlnL | s.e. | pBoot | ΔlnL | s.e. | pBoot |
| OUT(SUI(RUM(HIPPO,CET))) | ⟨−34823.0⟩ | | 0.951 | ⟨−30623.1⟩ | | 0.935 |
| OUT(SUI(CET(RUM,HIPPO))) | −42.2 | 24.2 | 0.045 | −48.3 | 22.9 | 0.016 |
| OUT((SUI,RUM)(HIPPO,CET)) | −59.6 | 22.5 | 0.004 | −34.9 | 20.0 | 0.040 |
| OUT(RUM(SUI(HIPPO,CET))) | −71.4 | 20.9 | 0.000 | −44.3 | 18.6 | 0.009 |
| OUT((SUI,CET)(HIPPO,RUM)) | −90.9 | 35.2 | 0.000 | −84.0 | 32.7 | 0.000 |
| OUT(SUI(HIPPO(RUM,CET))) | −68.9 | 21.0 | 0.000 | −67.4 | 20.4 | 0.000 |
| OUT(CET(RUM(HIPPO,SUI)))[a] | −170.0 | 33.0 | 0.000 | −153.0 | 29.1 | 0.000 |

[a]Tree according to the traditional classification.

which starts with a valine (GTG). No stop codon was identified for *COI*, and the stop codons of *COIII*, *NADH3* and *NADH4* are incomplete (TA or T). These features are consistent with observations made for several other mammalian mitochondrial genomes.

The control region of the sheep, a ruminant, contains tandemly organized repeats (Hiendleder *et al*. 1998), but like the whales (Arnason *et al*. 1991, 1993*b*; Arnason & Gullberg 1993), no such repeats occur in the control region of the hippopotamus. The origin of replication of the L-strand forms a hairpin loop with an 11 bp stem and a 13 nt loop in the hippopotamus, similar to that of other artiodactyls.

The phylogenetic tree shown in figure 1 has been reconstructed by QP–ML analysis of the aa alignment. It has been demonstrated in several previous phylogenetic analyses that the evolutionary rate of the mtDNA of the whales is considerably faster than that of the ruminants (Arnason *et al*. 1996, 1997; Xu *et al*. 1996*b*). As is evident in figure 1, the evolutionary rate of the hippopotamus is faster than that of the ruminants but slower than that of the whales. It seems likely therefore that the evolutionary rate of the ancestor of the hippopotamus–whale clade was accelerated relative to that of the ruminants, and that further acceleration took place in the cetacean lineage after the split between the hippopotamid and cetacean lineages.

The support values for relevant branches of the tree in figure 1 are given in table 1. The artiodactyl–cetacean clade (branch a, figure 1) received substantial support in the QP–ML and NJ analyses of both the aa and nt data sets, while the MP support for the same relationship was lower. The pig was basal in this clade, while the pecorans, cow and sheep, were identified as the sister group of the hippopotamid–cetacean clade (branch b, figure 1). The support for this relationship (branch c, figure 1) was strong in all analyses and data sets.

The log-likelihood and pBoot values were estimated for the 15 possible relationships between the artiodactyl–cetacean taxa included in figure 1, assuming a sister-group relationship between the cow and the sheep, and between the fin and the blue whales. The log-likelihood and pBoot values for the six best trees and the tree where both Artiodactyla and Suiformes are monophyletic, are given in table 2. The best tree was identical to that shown in figure 1. The second best tree joined the ruminants and the hippopotamus on a common branch to the exclusion of the whales. However, the log-likelihood value for this tree differed by more than one standard error from the log-likelihood value of the best tree, and the pBoot value was 0.016. The tree conforming to the traditional classification, monophyletic Cetacea and monophyletic Artiodactyla with the pig and the hippopotamus on a common branch, received least support among all possible trees.

## 4. DISCUSSION

The present findings which nest the Cetacea among recent artiodactyls, with strong support for a sister-group relationship between Hippopotamidae and Cetacea, suggest that both Artiodactyla and Suiformes are paraphyletic. Molecular data do not support the traditional suid–hippopotamid relationship (Beintema *et al*. 1986; Czelusniak *et al*. 1990), and most analyses of mitochondrial (Irwin & Arnason 1994; Arnason & Gullberg 1996) and nuclear (Sarich 1993; Gatesy *et al*. 1996; Gatesy 1997) sequence data, which have included suids, ruminants, hippopotamus and cetaceans, have identified a sister-group relationship between Hippopotamidae and Cetacea. On morphological grounds, the origin of Cetacea has been linked to Mesonychidae (Van Valen 1966), but the positioning of Cetacea within Artiodactyla challenges that notion. The now reported sister-group relationship between Hippopotamidae and Cetacea, however, may open other possibilities for examining the relationship between Cetacea and early Artiodactyla. The oldest hippopotamid fossils are of late Miocene age, but the relationship between Hippopotamidae and extinct artiodactyls has not been conclusively established, even though there is a considerable body of evidence that links Hippopotamidae with the semi-aquatic Anthracotheriidae (Colbert & Morales 1991). However, the position of the Anthracotheriidae in the artiodactyl tree has not been settled. The anthracotherids are traditionally

included in Suiformes, but the suiform characteristic of the anthracotherids are not pronounced, and as long as Hippopotamidae have been included in Suiformes the lack of distinct suiform characteristics in the Anthracotheriidae has been problematic. The present exclusion of Hippopotamidae from Suiformes permits other interpretations of artiodactyl–cetacean relationships, and with the establishment of the hippopotamid–cetacean relationship we find it reasonable to link the ancestry of Cetacea with the earliest Anthracotheriidae. Even though it might be coincidental, it is worth noting in this context that the most primitive anthracotheres had five digits in their forelimbs, the same number as occurring in Eocene Archaeoceti.

Molecular analyses of cetaceans have suggested an approximately contemporary diversification among the five extant cetacean lineages (Arnason & Gullberg 1996). In conjunction with the cetacean palaeontological record (Fordyce & Barnes 1994), this diversification has been dated to 32–34 million years before present (Ma BP; Arnason & Gullberg 1996). Extrapolation from this data set has suggested that cetaceans and ruminants, as represented by the cow, last had a common ancestor 60 Ma BP. After calibration for differences in evolutionary rates, the same approach to dating suggests that the ancestors of Hippopotamidae (probably Anthracotheriidae) and extant cetaceans last had a common ancestor *ca.* 54 Ma BP. This dating is compatible with the age, 52 Ma, of the oldest *Pakicetus* fossils (Gingerich *et al.* 1983), but is less compatible with the age, 54 Ma, of *Ambulocetus* fossils (Thewissen *et al.* 1994). The cetacean affinities of *Ambulocetus* have been questioned (Berta 1994), but the present estimate of the time of separation between the hippopotamus and the cetaceans does not give a conclusive answer to the question of these affinities. More precise estimates of the divergence times between ruminants, hippopotamids and cetaceans will become possible when a complete mtDNA of an odontocete species becomes available.

The present results have drawn attention to the problems associated with aligning morphological and molecular findings in a coherent scheme of systematics. This issue has been discussed in detail with particular emphasis on artiodactyl and cetacean relationships, even though the specific hippopotamid–cetacean relationship was unknown at that time (Graur & Higgins 1994). The establishment of a sister-group relationship between Hippopotamidae and Cetacea further underlines the problems associated with the assignment of systematic ranks to various groups of organisms. If the ordinal identity of Cetacea is maintained, then both the palaeontological and the molecular data suggest that Cetacea may be one of the very few eutherian orders originating after the K–T boundary.

We express our gratitude to Dr Eric Harley, Department of Chemical Pathology, University of Cape Town, South Africa, for the isolated hippopotamus mtDNA, to Dr S. Hiendleder for making the sheep mtDNA sequence available prior to publication and to Suzette Mouchaty and Kerryn Slack for proofreading. This study was supported by the Swedish National Sciences Research Council, the Jörgen Lindström Foundation, the Nilsson-Ehle Foundation, and by contract ERBCHR XCT 930254 from the European Commission.

## REFERENCES

Adachi, J. & Hasegawa, M. 1996*a* MOLPHY version 2.3: programs for molecular phylogenetics based on maximum likelihood. *Comput. Sci. Monogr.* **28**, 1–150.

Adachi, J. & Hasegawa, M. 1996*b* Model of amino acid substitution in proteins encoded by mitochondrial DNA. *J. Molec. Evol.* **42**, 459–468.

Anderson, S., de Brujin, M. H. L., Coulson, A. R., Eperon, I. C., Sanger, F. & Young, G. 1982 Complete sequence of bovine mitochondrial DNA. Conserved features of the mammalian mitochondrial genome. *J. Molec. Evol.* **156**, 683–717.

Arnason, U. & Gullberg, A. 1993 Comparison between the complete mtDNA sequences of the blue and the fin whale two species that can hybridize in nature. *J. Molec. Evol.* **37**, 312–322.

Arnason, U. & Gullberg, A. 1996 Cytochrome *b* nucleotide sequences and the identification of five primary lineages of extant cetaceans. *Molec. Biol. Evol.* **13**, 407–417.

Arnason, U. & Johnsson, E. 1992 The complete mitochondrial DNA sequence of the harbor seal *Phoca vitulina*. *J. Molec. Evol.* **34**, 493–505.

Arnason, U., Gullberg, A. & Widegren, B. 1991 The complete nucleotide sequence of the mitochondrial DNA of the fin whale *Balaenoptera physalus*. *J. Molec. Evol.* **33**, 556–568.

Arnason, U., Gullberg, A., Johnsson, E. & Ledje, C. 1993*a* The nucleotide sequence of the mitochondrial DNA molecule of the grey seal, *Halichoerus grypus*, and a comparison with mitochondrial sequences of other true seals. *J. Molec. Evol.* **37**, 323–330.

Arnason, U., Gullberg, A. & Widegren, B. 1993*b* Cetacean mitochondrial DNA control region: sequences of all extant baleen whales and two sperm whales species. *Molec. Biol. Evol.* **10**, 960–970.

Arnason, U., Gullberg, A., Janke, A. & Xu, X. 1996 Pattern and timing of evolutionary divergences among hominoids based on analyses of complete mtDNAs. *J. Molec. Evol.* **43**, 650–661.

Arnason, U., Gullberg, A. & Janke, A. 1997 Phylogenetic analyses of mitochondrial DNA suggest a sister group relationship between Xenarthra (Edentata) and ferungulates. *Molec. Biol. Evol.* **14**, 762–768.

Beintema, J. J., Fitch, W. M. & Carsana, A. 1986 Molecular evolution of pancreatic-type ribonucleases. *Molec. Biol. Evol.* **3**, 262–275.

Berta, A. 1994 What is a whale? *Science* **263**, 180–181.

Bibb, M. J., Van Etten, R. A., Wright, C. T., Walberg, M. W. & Clayton, D. A. 1981 Sequence and organization of mouse mitochondrial DNA. *Cell* **26**, 167–180.

Cao, Y., Adachi, J., Janke, A., Paabo, S. & Hasegawa, M. 1994 Phylogenetic relationships among eutherian orders estimated from inferred sequences of mitochondrial proteins: instability of a tree based on a single gene. *J. Molec. Evol.* **39**, 519–527.

Colbert, E. H. & Morales, M. 1991 *Evolution of the vertebrates*. New York: Wiley–Liss, Inc.

Czelusniak, J., Goodman, M., Koop, B. F., Tagle, D. A., Shoshani, J., Braunitzer, G., Kleinschmidt, T. K., de Jong, W. W. & Matsuda, G. 1990 Perspectives from amino acid and nucleotide sequences on cladistic relationships among higher taxa of Eutheria. In *Current mammalogy*, vol. 2 (ed. H. H. Genoways), pp. 545–572. New York: Plenum Press.

Dayhoff, M. O. 1978 *Atlas of protein sequence and structure*. Washington, DC: National Biomedical Research Foundation.

Felsenstein, J. 1991 PHYLIP: phylogeny inference package, version 3.52c. University of Washington, Seattle.

Fitch, W. M. 1971 Towards defining the course of evolution, minimum change from a specific tree topology. *Syst. Zool.* **20**, 406–415.

Fordyce, R. E. & Barnes, L. G. 1994 The evolutionary history of whales and dolphins. *A. Rev. Earth Planet. Sci.* **22**, 419–455.

Gatesy, J. 1997 More DNA support for a Cetacea–Hippopotamidae clade: the blood-clotting protein gene gamma-fibrinogen. *Molec. Biol. Evol.* **14**, 537–543.

Gatesy, J., Hayashi, C., Cronin, M. A. & Arctander, P. 1996 Evidence from milk casein genes that cetaceans are close relatives of hippopotamid artiodactyls. *Molec. Biol. Evol.* **13**, 954–963.

Gingerich, P. D., Wells, N. A., Russel, D. E. & Shah, S. M. I. 1983 Origin of whales in epicontinental remnant seas: new evidence from the early Eocene of Pakistan. *Science* **220**, 403–406.

Graur, D. & Higgins, D. G. 1994 Molecular evidence for the inclusion of cetaceans within the order Artiodactyla. *Molec. Biol. Evol.* **11**, 357–364.

Hasegawa, M., Kishino, H. & Yano, T. A. 1985 Dating of the human–ape splitting by a molecular clock of mitochondrial DNA. *J. Molec. Evol.* **22,** 160–174.

Hiendleder, S., Lewalski, H., Wassmuth, R. & Janke, A. 1998 Complete mitochondrial DNA sequence of a domestic sheep (*Ovis aries*) and comparative data from another major haplotype. *J. Molec. Evol*. **47**. (In the press.)

Irwin, D. M. & Arnason, U. 1994 Cytochrome *b* gene of marine mammals: phylogeny and evolution. *J. Mamm. Evol.* **2**, 37–55.

Irwin, D. M., Kocher, T. D. & Wilson, A. C. 1991 Evolution of the cytochrome *b* gene of mammals. *J. Molec. Evol.* **32**, 128–144.

Kishino, H. & Hasegawa, M. 1989 Evaluation of the maximum likelihood estimate of the evolutionary tree topologies from DNA sequence data, and the branching order in Hominoidea. *J. Molec. Evol*. **29**, 170–179.

Kishino, M., Miyata, T. & Hasegawa, M. 1990 Maximum likelihood inference of protein phylogeny and the origin of chloroplasts. *J. Molec. Evol*. **39**, 519–527.

Lopez, J. V., Cevario, S. O. & Brien, S. J. 1996 Complete nucleotide sequences of the domestic cat (*Felis catus*) mitochondrial genome and a transposed mtDNA tandem repeat (Numt) in the nuclear genome. *Genomics* **33**, 229–246.

Saitou, N. & Nei, M. 1987 The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Molec. Biol. Evol.* **4**, 406–425.

Sanger, F. 1981 Determination of nucleotide sequences in DNA. *Science* **214**, 1205–1210.

Sarich, V. M. 1993 Mammalian systematics: twenty-five years among their albumins and transferrins. In *Mammalian phylogeny* (ed. F. S. Szalay, M. J. Novacek & M. C. McKenna), pp. 103–114. New York: Springer.

Shimamura, M., Yasue, H., Ohshima, K., Abe, H., Kato, H., Kishiro, T., Goto, M. & Munechika, I. 1997 Molecular evidence from retroposons that whales form a clade within even-toed ungulates. *Nature* **388**, 666–670.

Strimmer, K. & von Haeseler, A. 1996 Quartet puzzling: a quartet maximum-likelihood method for reconstructing tree topologies. *Molec. Biol. Evol.* **13**, 964–969.

Tamura, K. & Nei, M. 1993 Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. *Molec. Biol. Evol.* **10**, 512–526.

Thewissen, J. G. M., Hussain, S. T. & Arif, M. 1994 Fossil evidence for the origin of aquatic locomotion in archaeocete whales. *Science* **263**, 210–212.

Ursing, B. M. & Arnason, U. 1998 The complete mitochondrial DNA sequence of the pig (*Sus scrofa*). *J. Molec. Evol.* **47**, 302–306.

Van Valen, L. 1966 Detratheridia, a new order of mammals. *Bull. Am. Mus. Nat. Hist.* **132**, 1–126.

Xu, X. & Arnason, U. 1994 The complete mitochondrial DNA sequence of the horse, *Equus caballus*: extensive heteroplasmy of the control region. *Gene* **148**, 357–362.

Xu, X. & Arnason, U. 1997 The complete mitochondrial DNA sequence of the white rhinoceros, *Ceratotherium simum*, and comparison with the mtDNA sequence of the Indian rhinoceros, *Rhinoceros unicornis*. *Molec. Phylogenet. Evol.* **7**, 189–194.

Xu, X., Gullberg, A. & Arnason, U. 1996*a* The complete mitochondrial DNA (mtDNA) of the donkey and mtDNA comparisons among four closely related mammalian species-pairs. *J. Molec. Evol.* **43**, 438–446.

Xu, X., Janke, A. & Arnason, U. 1996*b* The complete mitochondrial DNA sequence of the greater Indian rhinoceros, *Rhinoceros unicornis*, and the phylogenetic relationship among Carnivora, Perissodactyla, and Artiodactyla (+Cetacea). *Molec. Biol. Evol.* **13**, 1167–1173.